

Psychometric properties of the Chinese quality of life instrument (HK version) in Chinese and Western medicine primary care settings

Wendy Wong · Cindy Lo Kuen Lam ·
Kwok Fai Leung · Li Zhao

Accepted: 25 July 2011 / Published online: 10 August 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Background The Chinese Quality of Life Measure (ChQOL) had only been validated on a small number of selected subjects in Hong Kong and had never been tested in the Western medicine (WM) primary care setting.

Aims and objectives To test the psychometrics properties of ChQOL(HK version) in both TCM and WM general outpatient clinics.

Methods Three samples of Chinese adult patients [(1) 569 consulting TCM clinics for episodic illnesses; (2) 524 consulting WM clinics for episodic illnesses; (3) 205 consulting WM clinics for chronic disease follow-up] in Hong Kong were invited to complete the ChQOL(HK version) and the SF-36 Health Survey during their consultations and 2 weeks after consultations. The scaling assumptions, factor structure, convergent construct validity, reliability, responsiveness, and discriminatory power of the ChQOL were evaluated.

Results Majority of items satisfied the scaling assumptions. A two instead of 3-factor structure was found with physical form and emotion facets loading on one factor. Convergent construct validity was confirmed with moderate correlations with SF-36 scores. Internal consistency and test–retest reliability were satisfactory. The ChQOL(HK

version) was able to detect significant improvements 2 weeks after consultations, and it was able to discriminate between groups with different illness severity, age, and sex.

Conclusion The ChQOL(HK version) was shown to have satisfactory validity, reliability, discriminatory power, and responsiveness in both TCM and Western medicine primary care settings. The validity of the 3-domain scaling structure needs further evaluation.

Keywords Chinese quality of life instrument · Chinese medicine · Hong Kong · Primary care · SF-36

Introduction

A health-related quality of life (HRQOL) measure needs to be valid, reliable, and responsive to be useful as an evaluative tool in the clinical setting [1]. The Chinese Quality of Life Measure (ChQOL) was developed in Mainland China, based on the health concept of the Chinese culture to evaluate the effectiveness of Traditional Chinese Medicine (TCM) [2]. It has been validated and shown to be applicable to Chinese subjects in Mainland China and Hong Kong [2, 3]. The ChQOL was first adapted and pilot tested on 122 Cantonese speaking people in Hong Kong in 2007 [4]. The study supported the construct validity of the ChQOL (HK version) with a 3-domain structure and moderate correlation with the WHOQOL-BREF (HK) scores. It also showed good reliability with intra-class correlations between test–retest scores and Cronbach's alpha coefficients of internal consistency above 0.7 for all facets and domains. However, the generalizability of the results of this study is limited by a small and selective sample, and test–retest reliability was assessed within

W. Wong (✉) · C. L. K. Lam
Department of Family Medicine and Primary Care, LKS Faculty of Medicine, The University of Hong Kong, 3/F., Ap Lei Chau Clinic, 161 Ap Lei Chau Main Street, Ap Lei Chau, Hong Kong
e-mail: wwong110@hku.hk

K. F. Leung
Department of Occupation Therapy, The Queen Elizabeth Hospital, Kowloon, Hong Kong

L. Zhao
School of Chinese Medicine, The Chinese University of Hong Kong, Sha Tin, Hong Kong

2 days. Furthermore, cognitive debriefing on the content of the instrument was not carried out and responsiveness was not assessed.

Cognitive debriefing in our earlier study on content validity of this first Hong Kong version of the ChQOL found some linguistic and interpretation problems in 3 items that were subsequently revised [3]. We wanted to confirm the psychometric properties of the revised Hong Kong version of the ChQOL on a larger representative sample to confirm whether the instrument can be applied to the wider patient population in primary care in Hong Kong [4, 5].

TCM is mostly used as a complementary and alternative medicine (CAM) [5] in Hong Kong and western countries. More than 84.5% of Chinese population usually consult Western medicine (WM) as their main source of primary care [6], although many people in Hong Kong use both Chinese and western medicine. An important question is whether the ChQOL (HK version) is applicable to the evaluation of Western primary care services. The original ChQoL was intended to be used for the evaluation of the effectiveness of TCM and has never been evaluated among patients in the Western medicine (WM) setting. Theoretically the ChQOL is based on the health concept of the Chinese culture, so it should be applicable to all Chinese culture-based populations irrespective of the type of medicine they use.

Studies have shown significant differences in characteristics in patients consulting TCM from those consulting WM [6, 7]. The health care setting may also affect patients' interpretation and expectation of health. If it were proven to be valid and psychometrically adequate in both TCM and WM primary care, it can be used for the direct evaluation and comparison of integrated TCM and WM service in our primary care services.

Aim and objectives

Aim

The aim of this study was to determine the validity and other psychometric properties of the ChQOL(HK version) in both TCM and WM primary care settings in Hong Kong.

Objectives

The ChQOL(HK version) was tested in different clinical settings to establish as follow:

1. The psychometric properties of ChQOL(HK version) in patients consulting TCM primary care for episodic illness.

2. The psychometric properties of the ChQOL(HK version) in patients consulting WM primary care for episodic or chronic illness.

The relevance and spectrum of HRQOL domains may be different for patients with chronic and acute diseases, which may affect the psychometric performance of a HRQOL measure. Previous validation studies of the ChQOL were on patients with chronic diseases with no data from patients with episodic illness. Therefore, we tested the validity of the instrument separately on patients with episodic illness and chronic diseases in WM primary care.

Methods

Subjects

Three samples of Chinese adult (≥ 18) patients of either sex were recruited from one TCM and two WM general outpatient clinics from November 2005 to November 2007 in Hong Kong. The first was patients consulting the TCM outpatient clinic for an episodic of illness (sample 1, $n = 569$), the second was patients consulting either one of two WM outpatient clinics for an episodic illness (sample 2, $n = 524$). The third sample was patients attending one of two WM outpatient clinic (ALCC) for follow-up of their chronic diseases (sample 3, $n = 205$). An episodic illness is defined as "a new episode of illness for which the patient was consulting the clinic for the first time." A chronic disease is defined by the criteria of the U.S. National Center for Health Statistics [8], in the ALCGOPC were recruited from June to November, 2007. Table 1 reports the characteristics of the three samples.

The sample size estimated to detect an effect size change of 0.3 that corresponded to a minimum clinically important difference (MCID) in HRQOL [9, 10] was 278 patients with 80% power by paired *t* test (samples 1 and 2 for responsiveness). A target minimum of 400 patients was planned for samples 1 and 2 to allow for 30% drop-outs [11]. A minimum sample of 200 was planned for group 3 base on what is generally considered to be sufficient for factor analysis and psychometric testing [12].

Data collection

The aim, procedures, and prospective longitudinal nature of the study were explained, and written consent was obtained from each subject. Each subject completed the ChQOL(HK version), the Chinese (Hong Kong) SF-36 Health Survey, a structured questionnaire on socio-demography, morbidity, and health service utilization at

Table 1 Characteristics of patients from TCM and WM clinics

<i>n</i>	TCM episodic patients 569	WM episodic patients 524	WM chronic patients 205	HK general population (2009) 7,003,700
Age (mean/SD)	55.7 (14.3)	49.3 (15.9)	54.1 (11.4)	44.7 (16.2)*
Male, <i>n</i> (%)	189 (33.2)	213 (40.7)	95 (46.3)	889 (47.1)
Education level (%)				
Nil	10.9	5.8	9.8	22.9 ^a
Primary	28.6	21.5	37.6	16.1
Secondary	45.5	55.4	43.4	35.9
Tertiary	14.6	17.3	9.3	25.0
Unsure	0.4	0.0	0.0	0
Prevalence chronic disease (%)	45.3	33.8	83.9	NA
Hypertension	21.6	17.6	62.0	
Diabetes	4.0	1.1	12.7	
Cardiovascular disease	4.2	3.6	1.5	
Chronic pain	3.2	0.6	2.4	
Psychological disorder	0.7	0.4	0.3	
Others	66.2	76.7	21.2	
ChQOL(HK version) physical form (形) (mean/SD)	64.07 (15.42)	67.85 (13.47)	70.93 (12.90)	NA
Complexion	45.46 (18.41)	47.53 (17.03)	51.24 (15.29)	
Sleep	67.34 (26.77)	68.21 (22.60)	71.24 (21.43)	
Stamina	58.22 (23.39)	66.35 (20.19)	68.25 (18.89)	
Appetite and digestion	81.37 (19.77)	80.61 (15.94)	82.93 (14.22)	
Climate adaptation	67.94 (27.52)	76.53 (22.28)	81.26 (21.90)	
Vitality and spirit (神)	69.66 (18.50)	67.60 (15.47)	69.62 (15.00)	NA
Consciousness	79.80 (21.10)	82.39 (18.19)	87.40 (17.04)	
Thinking	59.02 (21.87)	58.71 (18.78)	57.61 (17.57)	
Spirit of eye	55.32 (32.29)	52.49 (25.33)	52.52 (25.01)	
Verbal expression	84.49 (18.66)	76.83 (17.92)	81.10 (18.42)	
Emotion (情志)	77.32 (17.74)	77.18 (15.34)	80.29 (14.60)	NA
Joy	67.00 (26.26)	61.17 (20.50)	63.66 (19.08)	
Anger	71.85 (22.93)	77.46 (20.24)	81.35 (20.19)	
Depress	81.60 (20.42)	83.38 (17.72)	86.36 (16.01)	
Fear	88.83 (19.69)	86.73 (19.04)	89.80 (17.94)	
Overall health	70.35 (14.49)	70.90 (12.32)	73.59 (12.40)	NA
SF-36				
Physical functioning (PF)	72.72 (24.09)	87.37 (16.06)	84.32 (18.05)	91.83 (12.89)
Role limitation due to physical problems (RP)	44.82 (43.14)	63.34 (39.15)	74.27 (38.10)	90.44 (17.93)
Bodily pain (BP)	49.16 (33.04)	63.15 (28.18)	72.33 (29.02)	83.98 (21.89)
General health (GH)	50.71 (26.65)	52.94 (21.74)	53.57 (22.83)	55.98 (20.18)
Vitality (VT)	58.03 (24.04)	60.97 (20.95)	66.08 (20.86)	59.92 (18.36)
Social functioning (SF)	83.28 (25.13)	85.04 (23.61)	89.46 (20.71)	91.19 (15.57)
Role limitation due to emotional problems (RE)	71.70 (41.51)	71.94 (39.73)	80.23 (35.11)	87.67 (18.16)
Mental health (MH)	74.89 (22.09)	74.31 (18.81)	79.67 (17.75)	71.46 (16.67)
Physical component summary (PCS)	33.46 (14.63)	42.76 (10.76)	43.38 (12.16)	50.00 (9.40)
Mental component summary (MCS)	55.55 (12.77)	52.13 (11.05)	55.62 (9.17)	50.00 (9.65)

NA no available information for comparison

* Significant difference between the four groups by ANOVA (* $P < 0.05$)

^a Education level was significantly different between four groups by Chi-square test ($P < 0.05$)

recruitment before the consultation with either the Chinese Medicine Practitioners (CMP) or the WM doctors. The CMP or WM doctors, blinded to the results of the interview, completed an evaluation on the severity of the subject's illness at the end of the consultation. All subjects were followed up 2 weeks after the initial consultation either by face-to-face or telephone interview. The same survey instruments and a global rating on change scale (GRS) were administered at the same time.

Study instruments

Chinese quality of life instrument

The ChQOL (HK version) and the Chinese (Hong Kong) SF-36 Health Survey were used to measure HRQOL in this study. The ChQOL (HK version) 2008 consists of 50 items in three specific domains: physical form (20 items), vitality and spirit (12 items), and emotion (18 items) [2–4]. The physical form domains contain facets of complexion, sleep, stamina, appetite and digestion, and climate adaptation. The vitality and spirit domains contain facets of consciousness, thinking, spirit of eye, and verbal expression. The emotion domains contain the facets of joy, anger, depress, and fear and anxiety. Each item is rated on a 5-point scale with items, facets, and domains score transformed to 0–100. Higher scores indicate better HRQOL.

The SF-36 health survey

The Chinese (Hong Kong) SF-36 Health Survey was adapted from the MOS 36-item Short-form Health Survey [13]. It is the most commonly used generic HRQOL measure consisting of 36 items measuring 8 domains: physical functioning (PF); role limitation due to physical problems (RP); bodily pain (BP); general health (GH); vitality (VT); social functioning (SF); role limitation due to emotional problems (RE); and mental health (MH). The domain scores can be summarized into two component summary scores, namely physical and mental component summary (PCS and MCS) scores. Each item is rated on a 2–5 point Likert scale, and the scale scores are transformed to a range from 0 to 100. The Chinese (Hong Kong) SF-36 Health Survey has shown to be reliable, valid, sensitive, and responsive in the local Chinese population [14, 15], and the acute version was used.

The global rating on change scale

The GRS asked the subjects to rate on the change in his/her own illness condition since the initial TCM/WM consultations. The response was given a score of zero for no change, +1, 2, or 3 for different degrees of improvement,

and –1, 2, and 3 for different degrees of deterioration [16,17].

Data analysis

The algorithms used in the original version ChQOL was followed in this study [2]. Floor and ceiling effects were considered significant if >15% of participants had a minimal or maximum baseline score [18], which might imply the scale is unable to detect an improvement or decline. Scaling assumptions were tested on item-facet, facet-domain correlations, and scaling success based on the hypothesized structure that consisted of 50 items, 13 facets, 3 domains, and 1 overall health dimension [19, 20]. All correlations were corrected for overlap [21].

Exploratory factor analysis was performed to test the hypothesized structure [2]. Principal components with eigenvalue >1 were extracted, and Varimax rotation was applied.

To evaluate concurrent construct validity, Spearman correlations between the domain scores of the ChQOL(HK version) and SF-36 scores were used. The ChQOL found that the physical form and vitality and spirit domain score correlated mostly with the GH and VT scores of the SF-36; emotion score correlated mostly with VT and MH scores of the SF-36. The overall health score correlated mostly with the GH and MH scores of the SF-36. Moderate to strong correlations (>0.3) were expected between corresponding domain scores. The changes in ChQOL scores were correlated with the changes in SF-36 scores and the GRS, based on the hypothesis that they should correlate whether they measure similar constructs of HRQOL.

Internal consistency was assessed as good when Cronbach's alpha >0.7 [22]. The test–retest reliability was assessed by intra-class correlation (ICC) between the ChQOL(HK version) scores of patients with chronic diseases, and in those reporting no change on GRS at 2 weeks interval. An ICC >0.75 is considered as excellent, 0.59–0.75 good, 0.40–0.58 fair, and <0.4 poor reliability [23]. The changes in the ChQOL(HK version) scores of patients, consulting either the TCM or WM clinics, for episodic illness were used to determine the responsiveness.

Responsiveness was measured by the Cohen's effect size [24], and change >0.3 was considered clinically significant [9, 10, 24]. The change in scores was further tested by Wilcoxon signed rank test, and *P* values <0.05 were considered statistically significant. The changes in ChQOL were correlated with the changes in SF-36 scores and the GRS to evaluate the validity of the change. The discriminatory power of the ChQOL(HK version) was tested by known group comparison of scores between patients with different levels of illness severity classified by CMP or WM doctors, and demographic groups. It was hypothesized

that patients with more severe illnesses, older or female would have lower ChQOL scores [15, 25]. All data analyses were carried out with the SPSS 17.0 version.

Results

Subjects

A total of 3,548 patients were approached and 1,735 subjects were excluded: (1) unable to communicate in Chinese, (2) too ill to complete the interview, and (3) age ≤ 18 . A total of 515 eligible subjects refused to participate resulting in a response rate of 71.6%. A total of 984 subjects were followed up either at the clinic or telephone with follow-up rate of 75.8%.

The characteristics of the 3 samples are shown and compared to the Hong Kong general population (2009) (Table 1). Subjects were generally older than the general population (mean age 56, 49, 54 vs. 45 years). The episodic health problems presented were mostly physical problems related to the respiratory system and musculoskeletal system (>60%). Hypertension, diabetes mellitus, and osteoarthritis accounted for >80% of the chronic diseases presented. Very few (<5%) patients had psychological problems. As expected, the SF-36 PCS and physical health-related scores (PF, RP, BP, GH, and VT) of our patient sample were much lower than those of the population norm. The mean SF-36 MCS score of the study subjects were higher than the general population mean, probably because of the older age of the study subjects as age has been found to be associated with better mental health-related quality of life [15].

Score distribution

All the items of the ChQOL(HK version) were answered with no missing data (Table 1). Patients with chronic illnesses generally had higher quality of life scores. There were few floor effects, only the facets complexion, stamina, and joy in which the proportions exceeded the standard 15% (Table 2). However, ceiling effects were found in all domains and most facet scores. WM chronic patients had the highest proportion of ceiling in most facets except those on thinking, spirit of eye, and joy; highest ceiling effects of the latter three facets were found in the TCM sample. Ceiling and floor effects were also found in several SF-36 scores especially in the RP and RE scales.

Scaling assumptions

The item-facet correlations between an item and its hypothesized facet score were >0.4 and similar in all

except 10 items (Table 3) in all patients groups. Most of these had correlations of greater than 0.3, which is acceptable. Item 6 of the facet of sleep (do dreams affect your quality of sleep?) had item-facet correlation of 0.35–0.37 in all three samples. In the facet of appetite and digestion, item 15 (do you often have digestion problem?), 16 (is your quantity of diet normal?), and 17 (do you have a good appetite?) had low correlations with the facet score (0.27–0.39) in both groups of WM patients. In the facet of climate adaptation, item 20 (do the changes of time in a day (e.g., day and night) cause any effect in your illness?) had weak item-facet correlations (0.24–0.25) in episodic patients of both TCM and WM clinics. The scaling success rate (higher correlation with hypothesized scale than others) at the facet levels were greater than 85%. The facet-domain correlations were all moderate to high with 100% scaling success in all three samples (Table 4). All facets scores also moderately correlated to the overall health score ($r = 0.45$ – 0.71).

Factor structure

Factor analysis on the item scores (data not shown) showed item 6, and item 20 had strong loadings (>0.7) on the hypothesized facets although they had low item-facet correlations; therefore, they were not excluded in the further factor analysis. Table 5 shows the results of exploratory factor analysis on the facet scores. Two factors were extracted, contrary from the three hypothesized for the three-domain structure of the original version of ChQOL [2]. Examination of the scree plots showed that the total variance was increased by a modest 7–10% by the inclusion of the next factor (Eigenvalue < 1), so it might not be appropriate to include this third factor. The physical form facets tended to merge with emotion facets in both WM patient groups but loaded on two factors in TCM patients. A one-factor structure was obtained when the domain scores loaded by patient groups, which matched the hypothesized structure for the physical form, vitality and spirit, and emotion domain scores to be summarized into the overall health dimension score.

Convergent construct validity

The correlations between the domain scores of the ChQOL (HK version) and the eight domain and two summary scores of the Chinese (HK) SF-36 Health Survey among the 3 patient groups are shown in Table 6. Scores of the physical form domain and vitality and spirit domains of the ChQOL (HK version) were most highly correlated with those of the GH and VT domains of the SF-36 ($r = 0.42$ – 0.64) while the score of the emotion domain was most highly correlated with the MH and RR domains of SF-36

Table 2 Floor and ceiling of ChQOL(HK version) and SF-36 Scores

<i>n</i>	% Floor			% Ceiling		
	TCM episodic patients 569	WM episodic patients 524	WM chronic patients 205	TCM episodic patients 569	WM episodic patients 524	WM chronic patients 205
ChQOL physical form	12.09	5.58	4.24	35.21	36.66	40.73
Complexion	16.52	12.22	6.83	6.11	1.48	2.27
Sleep	11.42	3.64	3.09	35.90	35.81	40.16
Stamina	19.92	7.25	7.32	27.96	36.56	39.76
Appetite and digestion	9.74	1.39	0.73	50.33	55.25	58.54
Climate adaptation	2.87	3.39	3.25	55.77	54.18	62.93
Vitality and spirit	7.34	5.60	5.21	42.85	26.50	35.81
Consciousness	10.19	1.04	0.65	49.17	50.61	69.92
Thinking	7.93	6.20	7.02	52.31	11.68	11.61
Spirit of eye	6.15	14.14	12.68	32.40	9.57	9.51
Verbal expression	5.10	1.04	0.49	37.52	34.13	52.20
Emotion	12.18	2.72	1.94	37.56	49.84	54.67
Joy	19.80	6.03	4.02	21.65	11.42	12.80
Anger	7.82	2.75	2.34	48.52	53.27	59.71
Depress	8.37	0.90	0.41	33.66	62.81	68.29
Fear and anxiety	12.71	1.20	0.98	46.40	71.86	77.89
Overall health	1.30	0.05	0.09	14.80	10.80	13.60
SF-36 (%floor/%ceiling of the general population)						
PF (0.20/46.0)	0.53	0.09	0.49	13.71	34.84	22.93
RP (0.60/64.40)	39.19	18.27	15.12	31.46	45.29	62.44
BP (0.50/54.70)	13.18	2.17	1.47	17.22	27.31	10.78
GH (1.00/0.50)	4.22	0.47	0.98	0.35	0.75	0.49
VT (0.20/1.70)	1.76	0.38	0.49	4.57	2.26	4.90
SF (0.10/70.80)	1.05	1.04	0.49	57.47	64.22	72.55
RE (0.30/55.40)	21.97	16.95	12.25	64.15	63.94	71.57
MH (0.00/4.50)	0.18	0.19	0.49	10.72	6.12	11.27

The % shown in brackets are the general population floor and ceiling proportions, respectively

PF physical functioning, RP role limitation due to physical problems, BP bodily pain, GH general health, VT vitality, SF social functioning, RE role limitation due to emotional problems, MH mental health, PCS physical component score, MCS mental component score

($r = 0.43\text{--}0.76$) in all the 3 patient samples. All ChQOL domains except emotion scores had moderate correlations with the SF-36 PSC, and the emotion domain score had moderate correlation with the SF-36 MCS.

Table 7 shows the correlations between the Chinese Medicine Practitioners (CMP) or doctors' rating on the severity of illnesses of patients and the HRQOL scores at presentation. A negative correlation was expected because higher severity score indicated more severe illness. There were significant but weak correlations between the CMP and Western medicine doctors' ratings and the ChQOL physical form score among episodic patients of TCM and WM clinics, but no significant correlation was found among chronic patients.

Table 8 shows the correlations between the change in HRQOL (ChQOL and SF-36) scores and GRS score. Only very weak correlations were found between subjects' GRS

score and the changes in ChQOL physical form or SF-36 scores in WM episodic patients. There were no significant correlations between the changes in ChQOL or SF-36 scores and GRS score except a very weak correlation between ChQOL physical form score and GRS in TCM patients.

Reliability

The Cronbach's alphas of all domains were greater than 0.8, and the overall health dimension were >0.9 in all groups (Table 9). All except those of three facets were >0.7 of ChQOL(HK version). The facets of climate adaptation and verbal expression had Cronbach's alpha below the optimal standard in all three groups. The 2-week test-retest reliability on the Intra-class correlation (ICC) coefficient of the ChQOL(HK version) among patients with

Table 3 Spearman item-facet correlations of the ChQOL(HK version)

Domains	Facets	TCM episodic patients		WM episodic patients		WM chronic patients	
		<i>n</i> 569		<i>n</i> 524		<i>n</i> 205	
		Correlations*	Scaling success rate [†] (%)	Correlations	Scaling success rate [†] (%)	Correlations	Scaling success rate [†] (%)
Physical form	Complexion	0.31–0.51	100	0.43–0.53	100	0.49–0.56	100
	Sleep	0.37–0.75	100	0.35–0.60	100	0.37–0.60	100
	Stamina	0.41–0.58	100	0.38–0.62	100	0.44–0.55	100
	Appetite and digestion	0.45–0.60	100	0.35–0.39	100	0.27–0.50	87.50
	Climate adaptation	0.25–0.47	100	0.24–0.51	100	0.42–0.53	100
Vitality and spirit	Consciousness	0.58–0.59	100	0.63–0.66	100	0.52–0.56	100
	Thinking	0.46–0.67	95	0.52–0.67	100	0.42–0.55	98.33
	Spirit of eye	0.54	100	0.48	100	0.54	100
	Verbal expression	0.37	91.67	0.43	100	0.42	91.67
Emotion	Joy	0.63–0.87	100	0.59–0.75	100	0.52–0.74	100
	Anger	0.40–0.67	96.67	0.42–0.65	100	0.48–0.64	100
	Depress	0.48–0.77	100	0.53–0.73	100	0.47–0.68	98.61
	Fear and anxiety	0.54–0.59	100	0.52–0.64	100	0.54–0.69	97.22

* Item-scale Spearman correlation between each item and its hypothesized facet score, corrected for overlap

† Scaling success means the item and hypothesized-scale correlation was higher than all item and competing-scale correlations. This rate was the proportion of total number of comparisons for all the items in each scale that were successful

Table 4 Spearman facet-domain correlations of the ChQOL(HK version)

Clinic	TCM episodic patients				WM episodic patients				WM chronic patients			
	<i>n</i> 569				<i>n</i> 524				<i>n</i> 205			
Domains	Physical form	Vitality and spirit	Emotion	Overall health	Physical form	Vitality and spirit	Emotion	Overall health	Physical form	Vitality and spirit	Emotion	Overall health
Facets												
Complexion	0.42* [†]	0.32*	0.35*	0.49*	0.42* [†]	0.36*	0.35*	0.53*	0.45* [†]	0.37*	0.38*	0.53*
Sleep	0.39* [†]	0.33*	0.43*	0.56*	0.45* [†]	0.31*	0.38*	0.56*	0.43* [†]	0.41*	0.42*	0.57*
Stamina	0.51* [†]	0.52*	0.47*	0.67*	0.57* [†]	0.49*	0.50*	0.71*	0.57* [†]	0.56*	0.50*	0.70*
Appetite and digestion	0.44* [†]	0.41*	0.44*	0.57*	0.51* [†]	0.31*	0.46*	0.56*	0.47* [†]	0.30*	0.45*	0.50*
Climate adaptation	0.37* [†]	0.32*	0.21*	0.46*	0.42* [†]	0.27*	0.33*	0.51*	0.43* [†]	0.40*	0.41*	0.57*
Consciousness	0.53*	0.64* [†]	0.47*	0.71*	0.43*	0.57* [†]	0.38*	0.63*	0.53*	0.59* [†]	0.52*	0.71*
Thinking	0.43*	0.67* [†]	0.37*	0.66*	0.41*	0.64* [†]	0.41*	0.66*	0.42*	0.57* [†]	0.36*	0.60*
Spirit of eye	0.49*	0.53* [†]	0.39*	0.70*	0.46*	0.53* [†]	0.36*	0.66*	0.46*	0.48* [†]	0.46*	0.67*
Verbal expression	0.34*	0.53* [†]	0.33*	0.55*	0.22*	0.45* [†]	0.25*	0.45*	0.38*	0.49* [†]	0.42*	0.59*
Joy	0.51*	0.47*	0.62* [†]	0.71*	0.43*	0.43*	0.46* [†]	0.63*	0.48*	0.46*	0.54* [†]	0.66*
Anger	0.40*	0.36*	0.58* [†]	0.60*	0.43*	0.31*	0.59* [†]	0.61*	0.44*	0.44*	0.59* [†]	0.64*
Depress	0.46*	0.43*	0.73* [†]	0.68*	0.49*	0.32*	0.69* [†]	0.65*	0.53*	0.42*	0.67* [†]	0.66*
Fear and anxiety	0.39*	0.31*	0.48* [†]	0.51*	0.48*	0.38*	0.58* [†]	0.64*	0.48*	0.44*	0.53* [†]	0.61*

* Spearman correlation between facets and domain scores was significant at $P < 0.01$

† Facet-domain Spearman correlation between each facets and its hypothesized domain score, corrected for overlap

no change on GRS ranged from 0.59 to 0.81, supporting the reliability of all scales. Similar results were found among patients with chronic diseases, further supporting the

reliability (Table 9). The lowest ICC was found in the complexion facet among both patients with chronic diseases (0.46) and those with no change on GRS (0.59).

Table 5 Exploratory factor analysis of facets and domains scores of the ChQOL(HK version)

Domains of ChQOL	Facets of ChQOL	TCM episodic patients (n = 569)			WM episodic patients (n = 524)			WM chronic patients (n = 205)		
		Factor 1 Factor loadings	Factor 2 49.74	% variance	Factor 1 Factor loadings	Factor 2 49.85	% variance	Factor 1 Factor loadings	Factor 2 52.16	% variance
Physical form	Complexion	0.48	0.23		0.42	0.37		0.39	0.37	
	Sleep	0.56	0.2		0.52	0.24		0.54	0.24	
	Stamina	0.41	0.59		0.58	0.45		0.53	0.54	
	Appetite and digestion	0.55	0.3		0.61	0.21		0.69	0.21	
	Climate adaptation	0.15	0.48		0.45	0.22		0.54	0.29	
Vitality and spirit	Consciousness	0.24	0.81		0.26	0.73		0.3	0.73	
	Thinking	0.16	0.83		0.25	0.78		0.13	0.85	
	Spirit of eye	0.31	0.61		0.25	0.70		0.37	0.56	
	Verbal expression	0.18	0.67		0.02	0.70		0.22	0.69	
Emotion	Joy	0.74	0.27		0.56	0.39		0.6	0.35	
	Anger	0.69	0.15		0.75	0.06		0.68	0.29	
	Depress	0.81	0.19		0.83	0.09		0.85	0.16	
	Fear and anxiety	0.66	0.18		0.75	0.15		0.72	0.19	
Dimension	Domain of ChQOL	Factor 1	% variance	Factor 1	% variance	Factor 1	% variance			
Overall health	Physical form	0.86	70.84	0.86	69.15	0.89	76.19			
	Emotion	0.83		0.83		0.88				
	Vitality and spirit	0.83		0.80		0.85				

Extraction method: principal component analysis; Eigenvalue > 1; Varimax rotation with Kaiser normalization

Table 6 Correlations between the domain scores of the ChQOL(HK version) and the Chinese (HK) SF-36 health survey

n	TCM episodic patients 569				WM episodic patients 524				WM chronic patients 205			
	Physical form	Vitality and spirit	Emotion	Overall health	Physical form	Vitality and spirit	Emotion	Overall health	Physical form	Vitality and spirit	Emotion	Overall health
Domain of SF-36												
PCS	0.40	0.38	0.16	0.37	0.49	0.31	0.32	0.45	0.56	0.45	0.36	0.53
MCS	0.46	0.35	0.66	0.58	0.42	0.34	0.63	0.56	0.42	0.34	0.6	0.52
PF	0.38	0.37	0.17	0.35	0.44	0.38	0.32	0.47	0.53	0.46	0.39	0.53
RP	0.42	0.34	0.3	0.41	0.40	0.23	0.36	0.39	0.43	0.36	0.38	0.43
BP	0.34	0.29	0.19	0.31	0.35	0.18	0.30	0.33	0.45	0.32	0.39	0.45
GH	0.64	0.48	0.48	0.63	0.58	0.42	0.49	0.60	0.62	0.59	0.51	0.68
VT	0.64	0.5	0.53	0.66	0.57	0.44	0.54	0.63	0.61	0.5	0.54	0.63
SF	0.37	0.24	0.34	0.39	0.34	0.18	0.41	0.37	0.35	0.19	0.35	0.33
RE	0.41	0.35	0.43	0.48	0.40	0.32	0.50	0.50	0.44	0.34	0.53	0.49
MH	0.55	0.43	0.76	0.72	0.45	0.38	0.68	0.61	0.53	0.48	0.72	0.66

All Spearman correlations are significant, $p < 0.01$

PCS physical component score, MCS mental component score, PF physical functioning, RP role limitation due to physical problems, BP bodily pain, GH general health, VT vitality, SF social functioning, RE role limitation due to emotional problems, MH mental health

Table 7 Spearman correlation between practitioner ratings and HRQOL scores

	Rating of Chinese medicine practitioners	Rating of Western medicine doctors	
	TCM episodic patients <i>n</i> = 569	WM episodic patients <i>n</i> = 524	WM chronic patients <i>n</i> = 205
ChQOL(HK version) physical form	−0.12*	−0.06	−0.07
Complexion	−0.06	−0.05	0.01
Sleep	−0.09*	−0.02	0.00
Stamina	−0.10*	−0.08*	−0.13
Appetite and digestion	−0.05	−0.03	−0.03
Climate adaptation	−0.11*	0.00	−0.09
Vitality and spirit	0.01	−0.13*	−0.10
Consciousness	0.00	−0.12*	−0.13
Thinking	0.04	−0.05	−0.05
Spirit of eye	−0.01	−0.10*	−0.08
Verbal expression	0.01	−0.10*	−0.03
Emotion	−0.03	−0.05	−0.06
Joy	−0.12*	−0.09*	−0.07
Anger	0.03	0.00	−0.03
Depress	0.03	−0.05	−0.03
Fear and anxiety	0.02	0.02	0.06
Overall health	−0.05	−0.09*	−0.08
Component summary score of SF-36			
PCS	−0.16*	−0.12*	−0.13
MCS	−0.13*	−0.02	−0.05
Domains of SF-36			
PF	−0.09*	−0.12*	−0.12
RP	−0.19*	−0.07*	−0.21*
BP	−0.11*	−0.11*	−0.03
GH	−0.20*	−0.09*	−0.15
VT	−0.15*	−0.06	−0.05
SF	−0.19*	−0.02	−0.21
RE	−0.10*	−0.04	−0.02
MH	−0.12*	−0.08*	−0.08

PCS physical component score, MCS mental component score, PF physical functioning, RP role limitation due to physical problems, BP bodily pain, GH general health, VT vitality, SF social functioning, RE role limitation due to emotional problems, MH mental health

* Spearman correlation is significant at the 0.05 level (* $P < 0.05$)

Responsiveness

Two weeks after the initial consultations, 829 TCM and WM episodic patients were followed up (Table 10). The ChQOL(HK version) and SF-36 scores of patients consulting with episodic illnesses significantly improved after

their consultations in both groups. Majority of the effect size changes of the ChQOL(HK version) exceeded the MCID standard of 0.3. Only the PCS and four domain scores of the SF-36 showed an effect size change of more than 0.3. The effect size changes of the SF-36 mental health-related (MCS, RE, and MH) scores were all less than 0.3, but the ChQOL emotion score had effect sizes ranged from 0.36 to 0.39 in the two patient groups.

Discriminatory power

Illness severity groups classified by the consulting practitioners, age groups, and genders were shown in Table 11. There was a trend for patients with mild diseases to have higher ChQOL scores than those with moderate or severe illnesses. The differences were statistically significant in one or more of the domain scores in episodic patients in either TCM or WM clinics. Moreover, the vitality and spirit domain of the ChQOL(HK version) was able to differentiate patients of different age groups showing poorer scores with increasing age. Female had lower scores than male in TCM episodic group. There was also the same trend in other patient groups although the differences did not reach statistical significance.

Discussions

This study included adult patients of all ages, both sexes, different educational levels, and a wide variety of health problems in different primary care settings. There was little difference in the psychometric properties of the ChQOL between patients with acute and chronic diseases or between TCM and WM, showing that the instrument is likely to be applicable all Chinese patients in primary care.

Score distribution

Ceiling effects were significant with a number of scales, which was not found in the previous studies [2, 4]. The discrepancy could be a reflection of relatively good health of a primary care population, as shown by ceiling effects in several SF-36 scales. Scales on sleep, appetite and digestion or consciousness might not be very useful for the evaluation of improvements in primary care. Adding more items that measure the extremes of quality of life might reduce floor and ceiling effects but this is beyond the scope of this paper. On the other hand, complexion, thinking, and spirit of eye are “bipolar” constructs that have more room for change in either direction. It should be pointed that the ChQOL overall health scale did not show any significant ceiling or floor effects, so it would be a useful “summary” scale for the evaluation on change in HRQOL in clinical trials.

Table 8 Spearman's correlations between change in ChQOL and SF-36 scores and global rating on change scale (GRS) score

<i>n</i>		TCM episodic patients	WM episodic patients	All patients					
		569	524	1,093					
		Patients' GRS at 2 weeks		Change in PCS	Change in MCS	Change in physical form	Change in spirit and vitality	Change in emotion	Change in overall health
SF-36	Change in PCS	0.07	0.16**		−0.35**	0.20**	0.15**	0.07	0.17**
	Change in MCS	−0.00	0.10**	−0.35**		0.28**	0.19**	0.42**	0.37**
ChQOL	Change in physical form	0.12*	0.12**	0.20**	0.28**		0.35**	0.40**	0.78**
	Change in spirit and vitality	0.00	0.05	0.149**	0.189**	0.35**		0.43**	0.77**
	Change in emotion	0.01	0.03	0.071	0.42**	0.40**	0.43**		
	Change in overall health	0.05	0.10**	0.169**	0.37**	0.73**	0.78**	0.77**	

Significant Spearman's correlation (* $P < 0.05$; ** $P < 0.01$)

Construct validity

A multi-method approach was used to evaluate the construct validity of the ChQOL(HK version) in terms of its scaling structure and correlations with other HRQOL measures and external criteria.

Scaling assumption

The majority of items satisfied the scaling assumptions in all patient groups. Correction for overlap was used in the item-scale correlation analysis, which is recommended because it is more stringent but may lead to lower correlations in scales that have few items [21, 26, 27]. Scores of items that define extreme conditions may be highly skewed in bipolar scales on sleep, appetite and digestion, complexion, and verbal expression [20], leading to relatively low item-scale correlations but they help to raise the “ceiling” or lower the “floor” of the scales, and improve the discriminatory power of the measure.

Items 20 (do the changes of time in a day (e.g., day and night) cause any effect in your illness?) and 16 (is you quantity of diet normal?) were most problematic in tests on scaling assumptions. Item 20 also had low content validity index on appropriateness (CVI) in previous content validation [3]. Further studies should be carried out on other Chinese people in Hong Kong to confirm whether these items were really non-homogenous. If proven, they should be eliminated from the HK version of the ChQOL.

Factor analysis

Factor analysis is commonly used to assess whether the items fall into the expected scales [28]. Items measuring the same concept are expected to be grouped by the same

factor (convergent validity) and vice versa for divergent validity. Confirmatory factor analysis (CFA) has been advocated by some to be method of choice for adaptation of HRQOL measures from one population to another [28, 29], but it might miss alternative factor structures. We used exploratory factor analysis (EFA) in this pilot study because the ChQOL is relatively new and its factor structure had only been replicated in a selected small sample of no more than 200.

The three domain scores loaded on one factor that explains 70% of the total variance, which supported the summation of the scores into a summary overall health score. Two instead of three factors were obtained from the facet scores loadings, contrary to the findings of the original study and an earlier study on Chinese people in Hong Kong [4]. The physical form and emotion facets tended to merge, and those of the vitality and spirit stood as a separate factor. The 2 factor structure was consistently found in all three primary care patient populations suggesting that they were likely to be true. It was noted in that both the physical form and emotion scales scores correlated more strongly with the SF-36 MCS than PCS score (Tables 6, 8) indicating that they both relate more with mental than physical health. The tendency to psychosomatization could be the reason for the merging of the physical and emotional factors. On the other hand, the possibility of differences in the functional meaning of the items between Hong Kong and Mainland Chinese need to be explored further [3].

It should be pointed out that results from exploratory factor analysis (EFA) can be difficult to interpret and there is no agreement on the best rotation method. Varimax rotation was used in this study to be consistent with the method used in the validation of the original ChQOL. We had carried out the factor analysis with oblique rotation (i.e., promax rotation) and found similar results of two

Table 9 Internal consistency and test–retest reliability of the ChQOL(HK version) by patient groups

<i>n</i>	TCM episodic patients	WM episodic patients	WM chronic patients			Patients with no change on GRS	
			ICC	95% CI	ICC	95% CI	
	569	524	205			736	
	Cronbach's alpha						
Domains/facets of ChQOL physical form	0.83	0.85	0.85	0.77	0.68–0.83	0.74	0.69–0.79
Complexion	0.64	0.70	0.71	0.46	0.26–0.61	0.59	0.49–0.66
Sleep	0.75	0.68	0.68	0.75	0.66–0.82	0.70	0.64–0.76
Stamina	0.76	0.78	0.76	0.71	0.60–0.79	0.75	0.70–0.80
Appetite and digestion	0.73	0.62	0.6	0.63	0.49–0.73	0.69	0.62–0.74
Climate adaptation	0.59	0.58	0.65	0.59	0.44–0.70	0.65	0.57–0.71
Vitality and spirit	0.87	0.86	0.84	0.65	0.52–0.75	0.76	0.71–0.80
Consciousness	0.76	0.78	0.70	0.61	0.47–0.72	0.71	0.65–0.76
Thinking	0.78	0.81	0.77	0.52	0.34–0.65	0.75	0.70–0.80
Spirit of eye	0.70	0.65	0.72	0.52	0.33–0.65	0.63	0.55–0.70
Verbal expression	0.57	0.59	0.54	0.52	0.35–0.65	0.62	0.54–0.69
Emotion	0.90	0.91	0.91	0.88	0.83–0.91	0.78	0.73–0.82
Joy	0.78	0.85	0.84	0.59	0.44–0.70	0.75	0.69–0.79
Anger	0.60	0.80	0.83	0.73	0.64–0.81	0.73	0.67–0.78
Depress	0.88	0.87	0.82	0.83	0.77–0.88	0.71	0.64–0.76
Fear and anxiety	0.74	0.76	0.79	0.82	0.75–0.87	0.63	0.55–0.70
Overall health	0.93	0.93	0.94	0.83	0.77–0.88	0.80	0.75–0.83
SF-36							
PF	0.86	0.83	0.81	0.81	0.73–0.86	0.80	0.76–0.84
RP	0.89	0.84	0.90	0.65	0.52–0.74	0.67	0.60–0.73
BP	0.88	0.88	0.85	0.81	0.74–0.86	0.67	0.60–0.73
GH	0.81	0.71	0.76	0.80	0.73–0.86	0.81	0.77–0.85
VT	0.80	0.79	0.77	0.77	0.69–0.83	0.71	0.64–0.76
SF	0.68	0.88	0.73	0.47	0.28–0.62	0.60	0.51–0.67
RE	0.91	0.86	0.86	0.64	0.51–0.74	0.50	0.39–0.59
MH	0.85	0.81	0.81	0.77	0.69–0.83	0.74	0.68–0.78
PCS	0.67	0.74	0.81	0.83	0.76–0.87	0.81	0.76–0.84
MCS	0.78	0.77	0.76	0.74	0.64–0.81	0.70	0.64–0.76

PF physical functioning, RP role limitation due to physical problems, BP bodily pain, GH general health, VT vitality, SF social functioning, RE role limitation due to emotional problems, MH mental health, PCS physical component score, MCS mental component score

factors. If further studies confirmed a two-domain structure, the validity of a shortened version of the ChQOL with elimination of non-discriminatory items and reorganization of the item-scale groupings should be explored.

Convergent construct validity with the SF-36 health survey

The ChQOL(HK version) scores moderately correlated with summary scores and most domain scores of the SF-36 confirming the two measures a similar construct. There were high (>0.6) correlations between all ChQOL(HK version) scores with the GH and VT scores of the SF-36 but relatively low correlations (<0.5) between the ChQOL physical form and SF-36 role functioning (PF, RP, RE and

SF) scores. These same results were found in the validation study of the original ChQOL [2]. The ChQOL, base on the TCM health concept, focuses mainly on general well-being and symptoms but not on role functioning. This may be a limitation when it is applied to patients with chronic diseases. Therefore, supplementation with a broader generic measure like the SF-36 or SF-12 may be needed. The correlations with SF-36 were mostly moderate indicating that the constructs of the two measures are related but not equivalent. The ChQOL captures Chinese culture-specific aspects of HRQOL by the unique facets of complexion, appetite and digestion, spirit of eyes, and verbal expression, which are not measured by any HRQOL measure developed in the West.

Table 10 Mean changes and effect size changes of HRQOL scores 2 weeks after TCM and WM consultations

<i>n</i>	TCM episodic patients 387* Mean change (effect size ^a)	WM episodic patients 442* Mean change (effect size ^a)
ChQOL (HK version)		
Physical form	9.17 (0.59)	-2.05 (-0.15)
Vitality and spirit	6.30 (0.34)	7.21 (0.47)
Emotion	6.42 (0.36)	5.97 (0.39)
Overall	7.35 (0.50)	3.76 (0.31)
SF-36		
PCS	6.74 (0.46)	5.02 (0.46)
MCS	1.52 (0.12)	2.35 (0.22)
PF	5.94 (0.25)	2.96 (0.19)
RP	18.28 (0.43)	14.65 (0.37)
BP	20.43 (0.62)	16.14 (0.56)
GH	6.13 (0.23)	7.68 (0.36)
VT	7.14 (0.30)	7.62 (0.37)
SF	5.72 (0.23)	6.17 (0.27)
RE	8.01 (0.20)	10.79 (0.28)
MH	5.22 (0.24)	3.67 (0.20)

All within group changes in scores after 2 weeks were statistically significant (* $P < 0.05$)

PF physical functioning, *RP* role limitation due to physical problems, *BP* bodily pain, *GH* general health, *VT* vitality, *SF* social functioning, *RE* role limitation due to emotional problems, *MH* mental health, *PCS* physical component score, *MCS* mental component score

^a Effect size was calculated as the difference between pre- and post-consultation scores divided by SD of baseline score

Only weak correlations were found between practitioners' rating and the ChQOL scores and were significant only among patients presenting with episodic illness. The same trend was found between SF-36 scores and practitioner ratings. The findings confirmed once again the difference in the constructs of subjective HRQOL perception and objective assessment (by the CMP/WM doctors). A self-limiting episodic illness may be judged by the doctor as a mild disease but it could be perceived to have significant impairment on quality of life by the patient. The reverse discrepancy in perception might occur with chronic diseases because patients are often asymptomatic even if the disease is quite severe. The results illustrated the importance of including both subjective and objective assessment in the evaluation of health outcomes.

Reliability

Internal consistency was demonstrated in patients using either TCM or WM primary care. The Cronbach's alphas of the overall health score was >0.9 , a standard that is

sufficient for not only group by individual evaluation. The reliability of the ChQOL(HK version) and SF-36 showed similar results that supported the application of both measures in both TCM or WM primary care.

The test–retest reliability of the ChQOL(HK version) was largely confirmed in this study. Facets of complexion, climate adaptation, thinking, spirit of eyes, and verbal expression had only fair reliability and were lower than those found by previous studies using a 2-day test–retest interval [2]. There is always a debate on the optimal interval for test–retest reliability [20, 30, 31]. A short interval like 2 days may inflate the test–retest reliability but the condition of patients could have changed with a longer interval (i.e., 2 weeks) [32]. A responsive HRQOL measure may show a low test–retest reliability because it can detect very small changes over time.

Responsiveness

Our study confirmed the responsiveness of the ChQOL in Hong Kong patients in both TCM and WM primary care settings. It was able to detect moderate effect size changes in HRQOL among patients after their consultations. Being Chinese culture specific, the ChQOL(HK version) emotion scale is more responsive than the SF-36 MCS and MH scales in detecting changes. The emotion domain in ChQOL (HK version) includes not only depression and anxiety but also joy and anger that are important emotions in the Chinese culture.

There was significant correlations between the changes in ChQOL and SF-36 scores supporting the validity that the former as a measure of HRQOL. The correlations between the changes in either HRQOL scores and GRS were very weak, suggesting that GRS might measure a different construct from HRQOL, or the “then recall” reference might be more susceptible to response shift.

Discriminatory power

The ChQOL(HK version) was able to discriminate different illness severity groups among patients presenting with episodic illnesses but not in those consulting for chronic diseases although the Chinese (HK) SF-36 PCS score was able to discriminate illness severity groups in all patient samples. This could be a reflection of a deficiency of the ChQOL. The ChQOL does not include the domains of role functioning and bodily pain, which are included in most other generic HRQOL measures. This illustrates that a culture-specific measure may not always be more sensitive than a cross-cultural measure. Further studies with different patient characteristics are needed to establish the discriminatory power of the ChQOL(HK version). Instead of choosing one over the other, the culture-specific ChQOL and cross-cultural HRQOL measures such as the SF-36

Table 11 HRQOL scores by severity of illness, age groups and genders

		Severity of illness		Age group			Genders	
		Mild	Moderate/severe	18–40	41–64	> 65	Female	Male
TCM episodic patients (<i>n</i> = 569)								
	<i>n</i>	133	436	92	301	176	380	189
Domains of ChQOL (HK version)	Physical form	66.8 (15.6)*	63.3 (15.2)*	62.1 (14.8)	63.9 (15.8)	65.3 (15.0)	61.8 (15.5)*	68.7 (14.1)*
	Vitality and spirit	69.8 (17.1)	69.7 (18.7)	73.9 (16.7)*	70.6 (18.4)*	65.7 (18.9)*	67.2 (18.2)*	74.7 (18.1)*
	Emotion	78.7 (17.2)	77.1 (17.7)	74.5 (18.0)	77.0 (18.3)	79.3 (16.6)	76.0 (18.6)*	80.0 (15.6)*
	Overall health	71.8 (13.7)	70.0 (14.5)	70.2 (13.8)	70.5 (15.0)	70.1 (14.0)	68.3 (14.6)*	74.5 (13.5)*
	SF-36	PCS	37.6 (13.4)*	32.0 (14.8)*	40.2 (13.0)*	34.4 (13.4)*	28.4 (15.8)*	31.4 (14.5)*
	MCS	57.7 (10.4)*	55.1 (13.2)*	51.3 (11.9)	54.4 (12.7)	59.7 (12.3)	55.4 (13.0)	55.8 (12.4)
WM episodic patients (<i>n</i> = 524)								
	<i>n</i>	351	173	157	267	100	309	215
Domains of ChQOL (HK version)	Physical form	68.7 (13.3)	67.6 (13.7)	67.3 (12.8)	68.6 (13.8)	68.8 (13.6)	65.6 (13.3)	72.1 (12.8)
	Vitality and spirit	69.2 (14.9)*	64.9 (15.6)*	71.7 (14.9)*	68.1 (14.9)*	61.22 (14.3)*	66.6 (15.2)	69.7 (15.0)
	Emotion	78.3 (15.2)	76.6 (15.8)	77.0 (15.2)	77.8 (15.3)	79.4 (15.3)	75.9 (16.0)	80.7 (13.5)
	Overall health	72.1 (12.1)*	69.8 (12.5)*	72.0 (12.1)	71.5 (12.3)	69.8 (11.8)	69.4 (12.2)	74.2 (11.6)
	SF-36	PCS	44.1 (10.2)*	41.5 (11.8)*	44.6 (10.5)*	42.9 (11.1)*	41.5 (10.4)*	42.0 (11.1)*
	MCS	52.5 (10.5)	51.5 (11.7)	50.8 (10.6)*	52.1 (10.8)*	54.9 (10.9)*	51.6 (11.4)	53.2 (10.0)
WM chronic patients (<i>n</i> = 205)								
	<i>n</i>	148	57	20	154	31	109	96
Domains of ChQOL (HK version)	Physical form	71.7 (12.6)	68.2 (13.6)	68.4 (13.4)	71.2 (12.7)	71.0 (14.0)	67.8 (13.3)	74.6(11.5)
	Vitality and spirit	70.4 (14.8)	67.4 (15.4)	68.0 (14.4)*	71.0 (14.0)*	63.2 (18.5)*	67.1 (15.1)	72.5 (14.5)
	Emotion	80.9 (14.2)	78.3 (16.9)	76.3 (13.9)	80.6 (14.5)	80.3 (16.3)	77.0 (16.4)	84.1 (11.1)
	Overall health	74.3 (12.0)	71.3 (13.6)	70.9 (11.1)	74.3 (12.1)	71.5 (14.4)	70.5 (13.1)	77.1 (10.6)
	SF-36	PCS	44.6 (11.8)*	39.9 (12.7)*	46.1 (15.7)*	44.0 (11.7)*	38.2 (11.8)*	40.9 (12.3)*
	MCS	56.0 (8.9)	54.4 (9.4)	51.4 (7.1)*	55.2 (9.1)*	59.0 (9.6)*	54.2 (10.3)	57.2 (7.4)

PCS physical component score, MCS mental component score

* Significant difference between age groups by Kruskal–Wallis *H* test (* *P* < 0.05)

Health Survey should complement each other since their HRQOL indicators are different.

Conclusion

The psychometric properties of the Chinese Quality of Life instruments (HK version) were shown to be satisfactory in both TCM and WM primary care settings. This study supports the application of the ChQOL (HK version) to Chinese patients using WM care. The results enable the ChQOL(HK version) to be used for direct comparison between the effectiveness of TCM and WM.

A few items of the ChQOL did not reach the expected standard in scaling assumptions, and some facets had

relatively low test–retest reliability. The 3-domain factor structure was not replicated in the ChQOL(HK version), which called for further studies to investigate the validity of a 2-domain structure. There is potential for the elimination of some items or revision of the scaling structure to produce a shorter version with improved psychometric properties.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Fitzpatrick, R., Fletcher, A., Gore, S., Jones, D., Spiegelhalter, D., et al. (1992). Quality of life measures in health care. I:

- application and issues in assessment. *British Medical Journal*, 305, 1074–1077.
2. Leung, K. F., Liu, F. B., Zhao, L., Fang, J. Q., Chan, K., et al. (2005). Development and validation of the Chinese quality of life instrument. *Health and quality of life Outcomes*, 3, 26.
 3. Wong, W., Lam, C. L. K., Leung, K. F., & Zhao, L. (2009). Is the content of the Chinese quality of life instrument (ChQOL) really valid in the context of traditional Chinese medicine in Hong Kong? *Complementary Therapies in Medicine*, 17(1), 29–36.
 4. Zhao, L., Leung, K. F., & Chan, K. (2007). The Chinese quality of life instrument: Reliability and validity of the Hong Kong Chinese version (ChQOL-HK). *The Hong Kong Practitioner*, 29, 220–232.
 5. Giovanni, A., Flavio, F., Maria, G. V., & Andrea, W. (2010). The ChQoL questionnaire: An Italian translation with preliminary psychometric results for female oncological patients. *Health and Quality of Life Outcomes*, 8, 106–133.
 6. Leung, G. M., Wong, I. O. L., Chan, W. S., Choi, S., & Lo, S. V. (2005). The ecology of health care in Hong Kong. *Social Science and Medicine*, 61, 577–590.
 7. Busato, A., Donges, A., Herren, S., Widmer, M., & Marian, F. (2006). Health status and health care utilisation of patients in complementary and conventional primary care in Switzerland—An observational study. *Family Practice*, 23, 116–124.
 8. MedicineNet.com. (2008). *Definition of chronic disease*. Available from: <http://www.medterms.com/script/main/art.asp?articlekey=33490>.
 9. Kazis, L. E., Anderson, J. J., & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, 27 (3, suppl), S178–S189.
 10. Wyrwich, K. W., Bullinger, M., Aaronson, N., Hays, R. D., Patrick, D. L., et al. (2005). Estimating clinically significant differences in quality of life outcomes. *Quality of Life Research*, 14, 285–295.
 11. Cowling, B. J. (2005). *Errors, power & sample size*. Department of Community Medicine, University of Hong Kong.
 12. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
 13. Ware, J. E., Snow, K. K., & Kosinski, M. (1995). *SF-36 health survey, manual & interpretation guide*. Boston, MA: The Health Institute.
 14. Lam, C. L. K., Gandek, B., Ren, X. S., & Chan, M. S. (1998). Test of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 Health Survey. *Journal of Clinical Epidemiology*, 51(11), 1139–1147.
 15. Lam, C. L. K., Launder, I. J., Lam, T. P., & Gandek, B. (1999). Population based norming of the Chinese (HK) version of the SF-36 health survey. *The Hong Kong Practitioner*, 21, 460–470.
 16. Jaescheke, R. J., Singer, J., & Guyatt, G. H. (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clinical Trials*, 10(4), 407–415.
 17. Osoba, D., Rodrigues, G., Myles, J., Zee, B., & Pater, J. (1998). Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*, 16(1), 139–144.
 18. McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: Are available health status surveys adequate? *Quality of Life Research*, 4, 293–307.
 19. Ware, J. E., Keller, S. D., Gandek, B., Brazier, J. E., & Sullivan, M. (1995). Evaluating translations of health status questionnaires: Methods from the IQOLA Project. *International Journal of Technology Assessment in Health Care*, 11, 525–551.
 20. Ware, J. E., & Barbara, G. (1998). Methods for testing data quality, scaling assumptions, and reliability: The IQOLA project approach. *Journal of Clinical Epidemiology*, 51(11), 945–952.
 21. Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
 22. Terwee, C. B., Bot, S. D., & De Boer, M. R. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60, 34–42.
 23. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlation uses in assessing rater reliability. *Psychological Bulletin*, 86, 20–428.
 24. Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
 25. Lam, C. L. K., & Lauder, I. J. (2000). The impact of chronic diseases on the health-related quality of life (HRQOL) of Chinese patients in primary care. *Family Practice*, 17(2), 159–166.
 26. Howard, K. I., & Forehand, G. G. (1962). A method for correcting error total correlations for the effect of relevant item inclusion. *Education Psychology Measurement*, 22(4), 731–735.
 27. Likert, R. (1932). A technique for the measurement of attitudes. *Arch Psychology*, 140, 5–55.
 28. Fayers, P. M., & Hand, D. J. (1997). Factor analysis, causal indicators and quality of life. *Quality of Life Research*, 6, 139–150.
 29. DeCoster, J. (1998). *Overview of factor analysis*. Retrieved from <http://www.stat-help.com/notes.html>.
 30. Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* (pp. 20–28). CA: Beverly Hills.
 31. Knapp, T. R. (1985). Validity, reliability, and neither. *Nursing Research*, 34(3), 189–192.
 32. Marz, R. G., Menezes, A., Horovitz, L., Jones, E. C., & Warren, R. F. (2003). A comparison of two time intervals for test-retest reliability of health status instruments. *Journal of Clinical Epidemiology*, 56, 730–735.